UNIVERSITÀ
DEGLI STUDI DI MILANO-BICOCCA

COURSE SYLLABUS

Statistical Models in Epidemiology (blended)

2122-2-F8203B008

Learning objectives

1. **Knowledge and Understanding:**

The course aims to introduce statistical methods to analyse (1) survival data: univariate estimation of survival times, group comparions, semi-parametric and parametric models, , competing risks, quantile regressionstudy designs; and and longitudinal data of repeated continuous and discrete measurements, conditional and marginal models (fixed and random effects), GEE models, missing data.

2. **Applying knowledge and understanding:**

Based on the methods taught in class, the students will apply the statistical models for survival and longitudinal data using Stata (StataCorp), widely used in the field of epidemiology and biostatistics. Also R and SAS will be briefly introduced.

3. **Making judgements:**

Having learnt methods and applications, the student will also show which methods and techniques are to be used given the study and the data at hand, by critically assessing the statistical and clinical implications of the models being used.

4. **Communication skills:**

The student will be able to explain how the methods can be applied and interpret the results in a clear and simple way.

5. **Learning skills:**

The students will learn how to analyse survival and longitudinal data from epidemiological and clinical studies, by understanding the methodology and by using Stata.

Contents

1. Survival Data Analysis: Introduction to Survival Data Analysis: Definitions. Censoring, Truncation, Survival Function, Hazard Function. Estimate and Comparison of Survival Curves. Cox Regression. Parametric Regression. Competing Risk. Quantile regression. Study design.

2. Introduction to longitudinal data: Summary Measure of Analysis of Longitudinal data. Modeling the mean response (analysis of response profiles, parametric and semi-parametric trends). Modeling the covariance. Linear mixed effects models. Generalized linear model for discrete longitudinal data. GEE models. Missing data in longitudinal studes.

Kleinbaum, D.G. and Klein, M. Survival Analysis, A self-learning text. (2013). Springer.

Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2013). Applied Longitudinal Analysis, Chapman & Hall CRC.

Jewell, N.P Statistics for Epidemiology. (2004). Chapman & Hall CRC.

Additional material is provided and made available on the course homepage

Detailed program

1. Survival Data Analysis: Introduction to Survival Data Analysis: Definitions and examples. Censoring and Truncation. Survival Function, Density Function and Hazard Function. Estimate and Comparison of Survival Curves (Log-rank test, Wilcoxon-Breslow-Gehan, Tarone-Ware test). Cox proportional Hazard Regression. Parametric Regression: the explonential and Weibull models. Competing Risk: net and crude survival, the cumulative incidence funzion, cause specific hazard and subhazard function. The Fine and Gray model. Quantile regression. Study design. Introduction to multiple events. Applications in R and SAS.

2. Introduction to longitudinal data: Summary Measure of Analysis of Longitudinal data. Modeling the mean response (analysis of response profiles, parametric and semi-parametric trends). Modeling the covariance structure (independence, exchangeable, unstructured, etc). Linear mixed effects models. Generalized linear model for discrete longitudinal data. Missing data in longitudinal studes.

Prerequisites

The course is offered in english, fluency in the english language is an important condition for successfully attending the lectures (listening comprehension and reading ability) and reading the additional material.

Teaching methods

Blended/elearning: frontal lectures, online lectures, group work, seminars, use of Stata. Online teaching will be used to allow students (not coming to class) to follow some of the lectures of the course. Lectures will be prepared and uploaded during the course. Moreover, homeworks will be discussed and possible solutions proposed once the homeworks are corrected.

Due to the current Covid-19 pandemic, teaching sessions will be broadcasted and recorded lectures will be made avalaible.

Assessment methods

1. Homeworks (30%), once a week
2. Final Written exam (70%)

3. Oral discussion (discretionary): it will serve to review the exam to better assess the student knowledge and unclear answers.

Both homeworks based on team work and the final exam will help to evaluate the student knowledge necessary to discuss the goal of the study being proposed to be analyzed using Stata, to run initial descriptive analyses, how to proceed with statistical inference and parameter interpretation, and overall model fitting.

The final score can be either increased or decreased after the oral review.

Students not willing or able to participate in the group work will be required to take the final exam which willl be given full weight.

Written exams are no different between attending and not attending students.

Due to the Covid-19 pandemic, exams will be run using online platforms (such as Zoom and WebEx), allowing for a public link.

Textbooks and Reading Materials

1. Kleinbaum, D.G. and Klein, M. Survival Analysis, A self-learning text. (2013). Springer.
2. Fitzmaurice, G. M., Laird, N. M., and Ware, J. H. (2013). Applied Longitudinal Analysis, Chapman & Hall CRC.
3. Jewell, N.P. Statistics for Epidemiology. (2004). Chapman & Hall CRC.

Additional material is provided and made available in the course homepage

Semester

II Semester, III

Teaching language

English